Homework 5

This assignment is due by **11:59pm on Friday 10/11** on gradescope. You should start a new page for each problem (if a problem has multiple parts, the parts can be completed on the same page). After completing the assignment, save your solutions as a single pdf and upload it to gradescope. Be sure to **assign your pages** on the gradescope upload page; failure to do so will result in a penalty.

You may collaborate with your peers on homework assignments and to discuss points of confusion with the instructor and/or TAs. However, you must write your own solutions independently and you may not look at or copy any one else's solutions. You must show all of your work to receive full credit (with the exception of row operations).

1. Use the $(7, 4)$ Hamming code defined in class with

$$H = \begin{bmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 1 \end{bmatrix} \text{ and } G = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \end{bmatrix}$$

to answer the following questions.

(a) Suppose Alice wishes to encode the message $M = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$. What is C_A?

$$C_A = GM = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}$$

(b) Suppose Bob receives the messages below. For each message, determine the original message M.

$$(i)\ C_B = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} \qquad (ii)\ C_B = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}$$

We must compute HC_B. If we get the zero vector, then $C_B = C_A$, in which case the first four entries of C_B is the original message M. If it is not the zero vector, then HC_B is one of the columns of H; if it is the ith column, then C_A is obtained from C_B by changing the i entry.

$\underline{(i)}$: $HC_B = \begin{bmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$. This is column two of H. Thus we must change the second entry of C_B. Thus $C_A = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}$ and so $M = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}$.

$\underline{(ii)}$: $HC_B = \begin{bmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$. Hence $C_B = C_A$ and so $M = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$,

2. In this problem, we will show that any message sent using the $(7, 4)$ Hamming code can be uniquely decoded. Define the *Hamming distance* between two vectors to be the number of components in which they differ. For example, the Hamming distance between $\begin{bmatrix}0\\1\\0\\0\\1\\1\\1\\0\end{bmatrix}$ and $\begin{bmatrix}1\\1\\1\\0\\1\\1\\0\\0\end{bmatrix}$ is three (since components 1, 3, and 7 are different). A *codeword* is a vector of the form $C_A = GM$ for some original message M (these are the vectors generated by the Hamming code; i.e. vectors in the column space of G).

   (a) Show that two codewords in the $(7, 4)$ Hamming code cannot be at Hamming distance 1 from each other. Hint: Use the following facts: if C_1 and C_2 are two codewords at distance 1 then for some i, $C_1 = C_2 + \mathbf{e}_i \pmod 2$; and codewords must be in the null space of H.

   Let C_1 and C_2 be codewords. Then they are both in the nullspace of H and so $HC_1 = HC_2 = \mathbf{0}$. Assume they have Hamming distance 1. Then $C_1 = C_2 + \mathbf{e}_i \pmod 2$ for some i and so $HC_1 = H(C_2 + \mathbf{e}_i) = HC_2 + H\mathbf{e}_i = \mathbf{0} + H\mathbf{e}_i = H\mathbf{e}_i$, which is the ith column of H. Since the columns of H are all nonzero, by definition, HC_1 is nonzero. But this contradicts the fact C_1 is in the nullspace of H. Hence C_1 and C_2 cannot be at Hamming distance 1.

   (b) Show that two codewords in the (7,4) Hamming code cannot be at Hamming distance 2 from each other. Hint: Use an argument similar to part (a).

   Let C_1 and C_2 be codewords. Then they are both in the nullspace of H and so $HC_1 = HC_2 = \mathbf{0}$. Assume they have Hamming distance 2. Then $C_1 = C_2 + \mathbf{e}_i + \mathbf{e}_j \pmod 2$ for some $i \neq j$ and so $HC_1 = H(C_2 + \mathbf{e}_i + \mathbf{e}_j) = HC_2 + H\mathbf{e}_i + H\mathbf{e}_j = \mathbf{0} + H\mathbf{e}_i + H\mathbf{e}_j = H\mathbf{e}_i + H\mathbf{e}_j$, which is the sum of the ith and jth columns of H. Since $HC_1 = \mathbf{0}$, we have that the sum of the ith and jth columns of H are 0. As we are performing operations $\pmod 2$, this implies that these two columns are identical. However, no two columns of H are the same, by construction (see H in problem 1). Thus we have reached a contradiction and so C_1 and C_2 cannot be at Hamming distance 2.

   (c) The above parts show that the Hamming distance between any two codewords must be at least 3. Explain why this implies that any message can be uniquely decoded.

Let $M_1 \neq M_2$ be two distinct messages. Then their codewords $C_1 \neq C_2$ are also distinct (since the first four entries of C_i are the entries of M_i). Let C_1' and C_2' be the transmitted messages (which each contain at most one error). Then since C_1 and C_2 have Hamming distance at least 3, C_1' and C_2' will have Hamming distance at least 1; hence $C_1' \neq C_2'$. So, no two codewords C_1 and C_2 can be confused for one another simply by looking at the transmitted messages.

3. (a) Let V and W be subspaces of $\mathbb{R}^n$. Show that if $V \perp W$, then $V \cap W = \{\mathbf{0}\}$.

Let $\mathbf{v} \in V \cap W$. Our goal is to show that $\mathbf{v} = \mathbf{0}$. Since $V \perp W$ and $\mathbf{v} \in V$ and $\mathbf{v} \in W$, we necessarily have that $\mathbf{v} \cdot \mathbf{v} = 0$. This is only possible if $\mathbf{v} = \mathbf{0}$. Thus $V \cap W = \{\mathbf{0}\}$.

(b) Use part (a) to deduce that any two subspaces of $\mathbb{R}^3$ represented by planes cannot be orthogonal.

Let V and W be subspaces represented by planes in $\mathbb{R}^3$. Since they both contain the origin, they cannot be parallel and thus $V \cap W$ must either be a line or a plane (if $V = W$). Hence $V \cap W \neq \{\mathbf{0}\}$ and so by part (a), V and W are not orthogonal.

4. Let $V = \left\{ \begin{bmatrix} 2x+y \\ x-y \\ 3x \\ y \end{bmatrix} \;\middle|\; x, y \in \mathbb{R} \right\}$.

(a) Interpret V as the columns space of a matrix A and use this to find a basis for V.

An arbitrary vector of V is of the form $\begin{bmatrix} 2x+y \\ x-y \\ 3x \\ y \end{bmatrix} = x \begin{bmatrix} 2 \\ 1 \\ 3 \\ 0 \end{bmatrix} + y \begin{bmatrix} 1 \\ -1 \\ 0 \\ 1 \end{bmatrix}$. Hence $V = \text{Span}\left\{ \begin{bmatrix} 2 \\ 1 \\ 3 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 0 \\ 1 \end{bmatrix} \right\}$ and so $V = C(A)$, where $A = \begin{bmatrix} 2 & 1 \\ 1 & -1 \\ 3 & 0 \\ 0 & 1 \end{bmatrix}$.

Since the columns of A are not multiplies of each other, they are linearly independent. Hence they form a basis for $C(A)$. That is, V has basis $\left\{ \begin{bmatrix} 2 \\ 1 \\ 3 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 0 \\ 1 \end{bmatrix} \right\}$.

(b) Use part (a) to determine $V^\perp$ and find a basis for $V^\perp$.

Since $V^\perp = (C(A))^\perp = N(A^T)$, we must find a basis for $N(A^T)$.
The RREF of A^T is $\begin{bmatrix} 1 & 0 & 1 & \frac{1}{3} \\ 0 & 1 & 1 & -\frac{2}{3} \end{bmatrix}$.
The system $A\mathbf{x} = \mathbf{0}$ is thus

$$x_1 + x_3 + \frac{1}{3}x_4 = 0$$
$$x_2 + x_3 - \frac{2}{3}x_4 = 0$$

Writing the system in terms of the free variables x_3 and x_4, we have the general solution

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} -x_3 - \frac{1}{3}x_4 \\ -x_3 + \frac{2}{3}x_4 \\ x_3 \\ x_4 \end{bmatrix} = x_3 \begin{bmatrix} -1 \\ -1 \\ 1 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -\frac{1}{3} \\ \frac{2}{3} \\ 0 \\ 1 \end{bmatrix}$$

Hence $\left\{ \begin{bmatrix} -1 \\ -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -\frac{1}{3} \\ \frac{2}{3} \\ 0 \\ 1 \end{bmatrix} \right\}$ is a basis for $V^\perp$ and

$$V^\perp = \text{Span}\left\{ \begin{bmatrix} -1 \\ -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -\frac{1}{3} \\ \frac{2}{3} \\ 0 \\ 1 \end{bmatrix} \right\}$$

(c) Verify your answer to part (b) is correct by taking all possible dot products of the basis vectors of V with the basis vectors of $V^\perp$.

$$\begin{bmatrix} 2 \\ 1 \\ 3 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} -1 \\ -1 \\ 1 \\ 0 \end{bmatrix} = 0 \qquad \begin{bmatrix} 2 \\ 1 \\ 3 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} -\frac{1}{3} \\ \frac{2}{3} \\ 0 \\ 1 \end{bmatrix} = 0 \qquad \begin{bmatrix} 1 \\ -1 \\ 0 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} -1 \\ -1 \\ 1 \\ 0 \end{bmatrix} = 0 \qquad \begin{bmatrix} 1 \\ -1 \\ 0 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} -\frac{1}{3} \\ \frac{2}{3} \\ 0 \\ 1 \end{bmatrix} = 0.$$

(d) What do V and $V^\perp$ look like geometrically? How does this compare to Problem 3b?

Since V and $V^\perp$ are 2-dimensional, they are planes in $\mathbb{R}^4$. Hence it is possible for planes to be orthogonal in $\mathbb{R}^4$, but not in $\mathbb{R}^3$ as seen in problem 3b.

5. The set of solutions to the linear equation $2x + y - 3z = 0$ is a plane W in $\mathbb{R}^3$ passing through the origin.

(a) Interpret W as the nullspace of a matrix A. What is A?

Since $2x + y - 3z = 0$ can be rewritten as $\begin{bmatrix} 2 & 1 & -3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \mathbf{0}$, $W = N(A)$, where $A = \begin{bmatrix} 2 & 1 & -3 \end{bmatrix}$.

(b) Find a basis for $W = N(A)$. To make your calculations easier in the next parts, choose a basis that does not contain fraction.

Since $a \neq 0$, the RREF of A is $\begin{bmatrix} 1 & \frac{1}{2} & -\frac{3}{2} \end{bmatrix}$. Hence the general solution is

$$\mathbf{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -\frac{1}{2}y + \frac{3}{2}z \\ y \\ z \end{bmatrix} = y \begin{bmatrix} -\frac{1}{2} \\ 1 \\ 0 \end{bmatrix} + z \begin{bmatrix} \frac{3}{2} \\ 0 \\ 1 \end{bmatrix}.$$

Hence $\left\{ \begin{bmatrix} -\frac{1}{2} \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} \frac{3}{2} \\ 0 \\ 1 \end{bmatrix} \right\}$ is a basis for W. Scaling the two vectors by 2 gives a different basis without factions, namely $\left\{ \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 3 \\ 0 \\ 2 \end{bmatrix} \right\}$.

(c) Find the projection of $\mathbf{b} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ onto W using the basis you found in part (b).

Let $A = \begin{bmatrix} -1 & 3 \\ 2 & 0 \\ 0 & 2 \end{bmatrix}$.

Then $A^T A = \begin{bmatrix} 5 & -3 \\ -3 & 13 \end{bmatrix}$ and $A^T \mathbf{b} = \begin{bmatrix} -1 \\ 3 \end{bmatrix}$.

To solve $A^T A \hat{\mathbf{x}} = A^T \mathbf{b}$, we find the RREF of $\left[A^T A \;\middle|\; A^T \mathbf{b} \right]$, which is

$$\left[\begin{array}{cc|c} 1 & 0 & -\frac{1}{14} \\ 0 & 1 & \frac{3}{14} \end{array} \right]$$

Hence $\hat{\mathbf{x}} = \begin{bmatrix} -\frac{1}{14} \\ \frac{3}{14} \end{bmatrix}$ and so $\text{proj}_W \mathbf{b} = A\hat{\mathbf{x}} = \begin{bmatrix} -1 & 3 \\ 2 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} -\frac{1}{14} \\ \frac{3}{14} \end{bmatrix} = \begin{bmatrix} \frac{5}{7} \\ -\frac{1}{7} \\ \frac{3}{7} \end{bmatrix}$.

(d) Let $T : \mathbb{R}^3 \to \mathbb{R}^3$ be the linear transformation given by $T(\mathbf{v}) = \text{proj}_W \mathbf{v}$. Find the projection matrix P associated to T.

The projection matrix is given by $P = A(A^TA)^{-1}A^T = \frac{1}{14}\begin{bmatrix} 10 & -2 & 6 \\ -2 & 13 & 3 \\ 6 & 3 & 5 \end{bmatrix}$

(e) Determine the rank of P without any row reduction; explain your reasoning.

Since $C(P) = \text{im}T = W$ and W is two-dimensional, the rank of P is 2.

6. Let A be an $m \times n$ matrix, and let $P = A(A^TA)^{-1}A^T$ be the projection matrix associated to the linear transformation that projects vectors in $\mathbb{R}^m$ onto the column space of A.

(a) Prove that $P^2 = P$.

$$\begin{aligned} P^2 &= (A(A^TA)^{-1}A^T)(A(A^TA)^{-1}A^T) \\ &= A(A^TA)^{-1}(A^TA)(A^TA)^{-1}A^T \\ &= A(I)(A^TA)^{-1}A^T \\ &= A(A^TA)^{-1}A^T \\ &= P \end{aligned}$$

(b) Interpret part (a) geometrically, in terms of projections. That is, explain why $P(P\mathbf{b})$ always equals $P\mathbf{b}$.

The product $P\mathbf{b}$ projects $\mathbf{b}$ onto the subspace $C(A)$ of $\mathbb{R}^m$; hence $P\mathbf{b} \in C(A)$. Now, $P(P\mathbf{b})$ is the projection of $P\mathbf{b}$ onto $C(A)$. But since $P\mathbf{b} \in C(A)$, we have that $P(P\mathbf{b}) = P\mathbf{b}$.

(c) What is wrong with the following "proof" that $P = I$? (Note that this statement is generally false!)

Proof:

$$\begin{aligned} P &= A(A^TA)^{-1}A^T \\ &= A((A^{-1}(A^T)^{-1})A^T \\ &= (AA^{-1})((A^T)^{-1}A^T) \\ &= (I)(I) \\ &= I \end{aligned}$$

□

Since A is not necessarily a square matrix, A^{-1} may not exist. Hence the simplification $(A^TA)^{-1} = A^{-1}(A^T)^{-1}$ used in the first step of the proof does not hold in

the general (note that the general property $(AB) = B^{-1}A^{-1}$ only holds if A and B are invertible).

7. Find the line of best fit through the data points $(1, 1)$, $(2, 1)$, and $(3, 0)$ using the method discuss.

   Let $y = mx + b$ be the line of best fit. To find it, we must find the least squares solution to the system

$$\begin{aligned} m + b &= 1 \\ 2m + b &= 1 \\ 3m + b &= 0 \end{aligned}$$

   This corresponds to the system $A\mathbf{x} = \mathbf{b}$, where $A = \begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \end{bmatrix}$ and $\mathbf{b} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$. The least squares solution is the solution to system $A^TA\hat{\mathbf{x}} = A^T\mathbf{b}$.

   Now, $A^TA = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \end{bmatrix} = \begin{bmatrix} 14 & 6 \\ 6 & 3 \end{bmatrix}$ and $A^T\mathbf{b} = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$. The RREF of $\left[\begin{array}{cc|c} 14 & 6 & 3 \\ 6 & 3 & 2 \end{array}\right]$ is $\left[\begin{array}{cc|c} 1 & 0 & -\frac{1}{2} \\ 0 & 1 & \frac{5}{3} \end{array}\right]$; hence the least squares solution is $\hat{\mathbf{x}} = \begin{bmatrix} -\frac{1}{2} \\ \frac{5}{3} \end{bmatrix}$ and so $y = -\frac{1}{2}x + \frac{5}{3}$ is the line of best fit.